QoS Techniques in Ad Hoc Networks
Michael Barry, Sean McGrath
Wireless Access Research Group
ECE Dept
University of Limerick, Limerick, Ireland
email: michael.barry@ul.ie, sean.mcgrath@ul.ie

Abstract
RRM algorithms are an essential part of end-to-end QoS provision in Ad Hoc networks. Judicious application of RRM algorithms, such as scheduling and admission control can be used to complement IP QoS strategies to provide an integrated QoS solution for Ad Hoc networks.

I. Introduction
IP Ad-Hoc Networks are quickly becoming part of today’s communications infrastructure. They enable new services and offer new means of providing services in public and private networks.

Quality of Service (QoS) in IP Ad Hoc networks remains an open issue. Due to the dynamic nature of the network, it is not possible to apply QoS Management techniques to negotiate Quality between users and networks. Quality enforcement mechanisms, on the other hand, can be applicable in ad hoc networks. IntServ and DiffServ techniques can be used to manage and control IP flows through queuing, marking and dropping of packets.

This paper outlines some of the difficulties faced in providing QoS in an IP mobile ad-hoc networking environment. Pure IP solutions are shown to be inadequate without support from the underlying radio technologies. A network model is constructed where different elements of Ad Hoc networks and their influences on RRM and QoS can be investigated. These elements include Mobility, Aggregation, Channel Sharing, Statelessness, and Routing. Each can be present to a greater or lesser extent in the network and will influence the RRM algorithms and overall QoS solution adopted.

II. QoS and Resource Management
IP Quality of Service concepts and architectures are displacing traditional notions of service and quality in existing ISDN, ATM and wireless networks. In traditional telecom networks Quality of Service is strictly managed and controlled, if not guaranteed, by operators and service providers. In general the IP QoS architecture can be viewed as consisting of two areas; service quality management and QoS enforcement techniques, both standardised by the IETF. QoS management is concerned with interaction with users, in the form of Service Level Agreements (SLAs) and SLA negotiation resulting in agreed performance expectations of the network [1]. QoS management is further concerned with the propagation of those expectations through the network, in the form of network-level and element-level policies.

In a mobile network, as well as being concerned with throughput, delay, call completion and other standard QoS requirements, QoS management must also be able to cater for user mobility and the vagaries of the radio channel. This may take the form of meta-policies, not related to any user per-se and so not negotiated in an SLA. Such meta-policies may include cell size and orientation, and handover acceptance metrics.

QoS enforcement techniques are the mechanisms by which traffic is controlled to conform to the QoS policies, it is through QoS enforcement that the network meets, or attempts to meet the performance expectations specified in the SLA and network and element policies. QoS enforcement, is in many ways another instrument of policy management, where the adopted enforcement techniques should best conform to customer expectations, network type and traffic characteristics. QoS enforcement techniques are also different in that they can be applied independently of any network-wide QoS management infrastructure, e.g. to police a particular bottleneck in the system [2].

QoS enforcement techniques are directly coupled with managing the available resources in a network. These resources may take the form of TDM timeslots, ATM cells, paths and channels in fixed networks and radio channels, codes, timeslots and frequencies in mobile networks. Enforcement techniques can be viewed as having two aspects, the resource management algorithms, and enforcement strategies. Management Algorithms include admission control, load balancing, and in wireless networks; handover, power control and code management.
Enforcement strategies derive meaningful parameters for resource management algorithms to work on. For example, Call Control protocols are used to derive a resource request that is used by the admission control algorithm. The prevalent QoS enforcement strategies in today's networks are DiffServ and IntServ in IP-based networks and ATM-based techniques in UMTS, and ATM backbone networks.

**QoS and Radio Resource Management**

Radio Resource Management (RRM), is another specialisation of Resource Management, with additional requirements for mobile and wireless systems. Radio Resource Management (RRM) is responsible for utilisation of the air interface resources to guarantee Quality of Service (QoS), to maintain the planned coverage area and to offer high capacity. RRM. As illustrated in figure 1 below, can be divided into a set of co-operative algorithms, handover, power control, admission control, load control, and packet scheduling.

Power control is needed to keep the interference levels in the air interface at a minimum and so is dependent on the physical layer of the wireless technology. Handovers are needed to handle the mobility of the user when they are moving from the coverage area of one cell to another. Admission control, load control, and packet scheduling are required to guarantee the quality of service and to maximise the system throughput with a mix of different bit rates, services, and quality requirements. As with resource management different enforcement strategies can be utilised to drive those algorithms with the aim of providing optimal performance applied network policies, and user-level QoS.

---

**Repositioning RRM for evolved networks**

Radio Resource Management is typically concerned with lower layer and MAC issues for the control of local radio resources in wireless networks. This has provided efficient use of the radio spectrum when the available services are well known and limited, such as voice. Likewise in the overall context of Network Resource Control and Management, Radio Resource Management strategies are often treated as a closed area, independent of the overall resource management issues and so the scope and visibility of RRM is limited to within the Access Network.

The introduction of new types of data services, the potential of new wireless network types including ad-hoc networks and the advent of all-IP wireless networks require a new perspective on Radio Resource Management and how it is integrated into overall Network Resource Control. Figure 2 below positions Radio Resource Management in relation to other wireless system aspects. RRM is the bridge between a range of wireless technologies and higher layers and services.

---

**III. QoS in Ad Hoc Networks**

Applying QoS management architectures in an ad-hoc IP network is impractical, as these typically require a level of centralisation that is not present in such networks. Likewise SLA negotiation and policy deployment and enforcement do not make sense without some administrative domain and architectures to support them.

While IntServ ensures that resources have been reserved at the IP layer, this will not necessarily be reflected at the radio level if there is no support for QoS in the MAC as is the case for IEEE 802.11. Neighbouring nodes can access the channel and transmit data using standard IEEE 802.11 mechanisms without consideration of the QoS requirements of surrounding nodes. If the radio environment is congested then the IP Reservation,
and the dependent QoS assurances cannot be met [4].

Figure 3: Interference and Reservations in ad-hoc nets

One solution to this is for each node to periodically broadcast its required reservations, which are then respected by all surrounding nodes and not only the nodes through which the IP flow is routed. Assuming that a single channel is used in the network, if source and destination nodes are very separated in distance then the entire network may be constrained by reservations for individual IP flows. This drastically curtails the system throughput, effectively treating the entire network as a single channel, rather that a set of interacting nodes separated by time and distance.

DiffServ techniques can also be used in IP ad-hoc networks. As with the IntServ techniques outlined above if there is no QoS support in the lower layers, then these are of limited use. Application of a single policy across an entire network may result in improved QoS for some service types. The main advantage of this scenario is that it reduces signalling overhead on the air interface. Furthermore by policing traffic on a service aggregate rather than on a per node basis it is possible to provide greater service differentiation between priority services and non-priority services, regardless of originating mobile.

The disadvantage of a pure DiffServ scenario is that for excess traffic to be present at a node that traffic must have been transmitted using radio resources. As the excess traffic is to be dropped then those radio resources have effectively been wasted. Again, non-adaptive or misbehaving nodes or applications may generate a large amount of excess traffic wasting precious radio resources and having a detrimental effect on quality in the system.

IV. Resource Management in Ad Hoc Networks

From the point of view of the upper layers, the IP and Session layers, Radio Resource Management provides an essential part of QoS, Mobility and Routing Functions. The main function of Radio Resource Management at this level is to allocate and control radio resources amongst the different users and terminals in the environment according to their individual service needs and priorities. Resources may be allocated deterministically as in 3G/UMTS or statistically as in WLAN DCF. This provides a platform for providing a network path for the transport of messages to the terminal.

Through manipulation of the Radio Resources, guarantees of bandwidth and delay, per service type or per session can be provided. This allows the integration of the radio link, with predictable behaviour, into the overall end-to-end QoS provision for the service type or session. Different implementation strategies for RRM integrate with different QoS techniques, including DiffServ and IntServ. Different levels of signalling can also be used in implementing RRM from intensive ATM-like signalling to none. If no signalling is used then some implicit means of allocating Radio Resources is required.

User and service mobility is supported through handover control in RRM. During a handover, not only must a new connection be established for the routing and transport of messages, but that connection should also conform to any bandwidth and delay guarantees previously allocated to the service type or session to facilitate continuity in end-to-end QoS as perceived by the end user.

Five Properties of Ad Hoc Networks

Rather than directly applying existing IP QoS enforcement to Ad Hoc networks it is first necessary to characterise Ad Hoc networks in such a way that different QoS enforcement strategies and their impact on available, or potential RRM algorithms can be identified and analysed. In fact it can be said that applying enforcement strategies without regard to their impact on RRM algorithms, as in the above, section provides no significant gains in QoS whatsoever. Mobility, aggregation, channel type, flow state, and routing are properties that are common, to some degree, in all ad hoc networks.

As users move within the network the traffic transiting the network will fluctuate accordingly. Variations in the traffic stream can reflect increases or decreases in user activity, or reflect that users are moving in and out of the immediate radio environment. Implications are on load balancing (as new users enter the system can the system return to stability) and Admission Control (under what circumstances should users be allowed into the system).

Aggregation is concerned with the grouping of data flows within the system. Flows can be treated
individually based on different n-tuples e.g. (source, destination, port, type) or aggregated together, according to destination, type, etc. The trade off is in the load balancing/scheduling, it can be easier to manage a small number of aggregate flows, as opposed to a large number of individual flows. Aggregation can also affect admission control and mobility – individual flows may still need to be admitted to an aggregated set.

If a network maintains state information related to each flow in the network, and its QoS, characteristics then that network can be said to be stateful. If no state information is maintained then the network is stateless. If a system utilises state information some mechanism must exist to set it up and maintain that information. Some RRM decisions, e.g. admission, can be made with respect to current conditions and may not reflect the actual needs of admitted traffic. Likewise load balancing may occur by traffic type (voice vs. data) or packet type (TCP vs. UDP) without regard to individual flow state.

Channels can be shared, with a number of nodes competing to access the channel simultaneously as in IEEE 802.11, or distinct, with specific bandwidth allocated to individual nodes or services. Systems such as GPRS/UMTS adopt a hybrid approach with some channels reserved for simplex or duplex communication for a single node and some channels being shared amongst a number of nodes.

Routing, while not typically related to QoS and RRM is however an unavoidable fact in IP Ad Hoc networks. The goal of routing in Ad Hoc networks is to locate and propagate paths for traffic flow in the network. Ideally, this routing information should also convey information on the usefulness of the route for QoS purposes. Route metrics such as link delay, available bandwidth and number of hops can be used to select usable links in the network. They can also be useful in Admission control and mobility, where an incoming flow may be rejected because no suitable route to the destination exists.

V Selection and simulation of RRM algorithms

While DiffServ itself is inadequate for QoS in Ad Hoc Networks, Differentiated Service mechanisms that utilise statelessness and aggregation can be used to provide long timescale service provision in mobile and Ad Hoc environments.

Providing differentiated services in a mobile environment requires that the radio some degree of separation between different types of services. This separation can be based on the DiffServ field in IP packets [5]. DiffServ oriented RRM has to ensure that available radio resources are shared among active users, while at the same time ensuring that different traffic types receive service in a differentiated manner.

RRM must be adaptive and robust to both internal and external dynamics. Load control and scheduling must offers effective protection for the differentiated traffic classes against traffic fluctuations in lower classes. These algorithms should be robust to changes in the external environment; for example, fluctuations of traffic due to mobility must have a predictable and limited effect on the delay and loss experienced by all service classes in the network.

Decentralized and adaptive mechanisms can more efficiently solve these problems in a distributed ad hoc environment in comparison to centralized ones. Distributed control of the radio resources may result in more productive use of radio resources ratio. Distributing control of the radio resources allows mobile hosts within the same class to compete for radio resources and achieve acceptable fairness, while at the same time offering differentiated access to different service classes.

A modified IEEE 802.11 [7] is used to provide load balancing and service differentiation for real time and non-real time traffic. By manipulating CW values it is possible to constrain the backoff mechanism of IEEE802.11 [6] which control how nodes access the transmission medium. Utilising multiple backoff schema for services in different nodes provides a distributed scheduling and load control mechanism in the system.

The traffic mix consists of mobile hosts sending high priority voice or and video and hosts starting best effort greedy TCP connections. Voice traffic is modeled using an on/off source with exponentially distributed on and off periods of 300ms average each. Video traffic is generated using a 65kbps constant rate source. Traffic is generated during the on periods at a rate of 32 kbps with a packet size of 160 bytes, thus, the inter-packet time is 40 ms. During all simulations the channel rate is 2 Mbps. Simulations were performed using the ns-2 network simulator developed by the VINT Project [10] with the wireless extension produced by the MONARCH Group [9].

During simulation, the channel load is increased by adding a new voice, video (64kbps constant rate source) and TCP session periodically every 5 seconds. The voice and video sources use CWmin
and $CW_{max}$ values of 16 and 64, while the TCP traffic uses 128 and 1024, respectively.

Figure 4: Aggregate Delay for increasing traffic

Figure 4 above shows the delay throughout the simulation for the three traffic types. It can be observed that the delay increases for all service types but the delay separation is efficiently maintained from low load up until the channel is saturated.

For best-effort traffic the achievable throughput is of more importance than delay. Fig. 3 shows that the distributed load control algorithms enable the best-effort adaptive TCP traffic to utilize any free capacity unused by high priority sources.

It can be observed that even at the saturation point, the TCP traffic is not completely starved. This is due to the statistical and non-deterministic nature of the IEEE 802.11 CSMA/CA [11].

Figure 5: Aggregate throughput for increasing traffic

Admission Control

To support real-time services it is not sufficient to ensure that high priority traffic gets better service than best effort. In most cases applications require absolute and not relative service quality, (e.g., voice or video). If a mobile host realizes that the channel is unable to meet its delay and loss requirements, it can refrain from loading the channel further or reduce its application traffic demands, (e.g., by increasing compression).

The channel can be in one of three states from the perspective of a new traffic flow [8]: throughput limited, delay limited, or not congested. The "not congested" state means that the channel can probably serve the new traffic flow without severely degrading the channel state. Throughput limited state means that the channel is highly congested, and the new traffic flow would not be able to achieve the required throughput. The delay limited state means that although the channel may not be fully utilized, the delay experienced by the traffic source would probably exceed the requirements.

Admission control algorithms determine whether the channel can support a new traffic stream or not and admit or reject a new session accordingly.

The natural location of the AC algorithm in cellular systems is at the base station that the mobile host is associated with. If all the base stations in the area execute the same algorithm, a globally stable state can be maintained. This is true even if cell areas overlap and share radio resources. In an Ad Hoc network where channel properties may vary amongst different nodes it is preferable that all hosts also participate in Admission to ensure that the service quality will be met for the new stream.

Admission is granted if algorithms at the receiving and sending nodes (and all nodes in between) admit the new request.

Ten receiving are placed randomly on a 400m by 400m with their coverage areas significantly overlapping. One hundred mobile hosts were placed randomly in the coverage area. Half of the mobile hosts randomly generate Web sessions and the other half randomly generate voice traffic. Every sender is associated with the nearest receiver. The length of the voice sessions and the inter-arrival times between connection requests were exponentially distributed. The average session length was 30s. Upon completion of a session, a mobile host attempted a new call after an average waiting period of 10s.

Admission control was applied to delay sensitive voice sessions. When the estimated delay exceeds 10 ms, new voice sessions were rejected from service. If
accepted, the voice packets use the modified MAC algorithm with \(cw_{\text{min}} = 32\) slots and \(cw_{\text{max}} = 64\) slots, while the Web sessions use values 64 and 1024, respectively. There was no admission control applied to Web traffic. Fig. 12 shows the average per-packet delay of traffic in the entire coverage area.

![Figure 6: Per packet delay for admitted voice calls](image)

**VI. Conclusions**

The advent of heterogeneous data and voice networks, supporting new types of network and radio technologies, calls for a new approach to Radio Resource Management. Rather than being radio and layer-2 oriented, RRM provides an essential part of any end-to-end QoS solutions, complementing existing IP oriented strategies and services. This new vision of RRM is of utmost relevance in Ad Hoc networks where traditional centralised approached to RRM are not possible.

The primary characteristics of Ad Hoc networks: mobility, aggregation, statelessness, routing and shared channels. Any QoS solution adopted must accommodate these characteristics to a greater or lesser degree. RRM is the building such solutions by enabling distributed scheduling, admission and mobility algorithms, providing reliable levels of services across a range of traffic types.

**VII. Acknowledgements**

This work is conducted as part of the REALM project funded by Enterprise Ireland. Ericsson Ireland has supported Michael Barry’s work on this project.

**VIII. References**


[9] CMU MONARCH Project, ww.monarch.cs.cmu.edu
